

Japanese Patent Laid-open Publication No. HEI 11-283036 A

Publication date : October 15, 1999

Applicant : Toshiba Tec K. K.

Title : SUBJECT DETECTION DEVICE AND SUBJECT DETECTION METHOD

5

(57) [Abstract]

[Object] A detection of a subject is carried out while taking into consideration the individual difference in a feature area position of a discrimination subject, variations in size and variations in pickup environments, so as to improve the detection rate, and the necessity of verification of the positional relationship between feature areas is eliminated so that a simple process is realized.

[Means to Solve the Problems] An image including a face is inputted from an image input section 1, and is stored as a gray-scale image, and this gray-scale image is subjected to a differential process by an image processing section 2 to generate an edge image, and this is then stored. Next, at a position detection section 3, with respect to the gray-scale image and the edge image, by using area models which are a plurality of preset judgment element acquiring areas corresponding to featuring areas such as the eyes, nose, mouth, etc. of a face image, each of the area models is inserted to the image while the position being successively specified, and the amount of feature is extracted from each of the judgment

element acquiring areas so that based upon the extracted amount of feature, a face image is detected from the image.

[0016]

5 [EMBODIMENTS OF THE INVENTION] Referring to Figures, the following description will discuss embodiments of the present invention.

(FIRST EMBODIMENT) This embodiment exemplifies a case in which a face detecting operation is carried out with respect  
10 to the discrimination subject detection.

[0017] Fig. 1 is a block diagram that shows the entire structure of a subject detection device which is provided with an image input section 1 serving as an image input means for inputting an image including a person's face, an image  
15 processing section 2 for carrying out predetermined processes on the image inputted by the image input section 1, and a position detection section 3 which, based upon image information inputted by the image input section 1 and the image information processed by the image processing section 2, detects the  
20 position of the face image.

[0018] As illustrated in Fig. 2, the image input section 1 is provided with a CCD camera 11 for picking up a person's face and for outputting digital gray-scale image information including the face, an image input board 12 for taking in the  
25 digital gray-scale image information from the CCD camera 11,

and an image memory 13 for storing the digital gray-scale image information taken in by the image input board 12. Here, the image to be inputted may be a color image.

[0019] The above-mentioned image processing section 2 reads  
5 the digital gray-scale image information from the image memory 13, subjects this to a differential process so as to be converted to edge image information, and stores the resulting information in the image memory 13. In other words, the gray-scale image, as illustrated in Fig. 3(a), is subjected to the differential  
10 process to generate the edge image as shown in Fig. 3(b), and this is stored in the image memory 13. With respect to the extraction of the edge, a Laplacian filter or a Sobel filter, etc., which has been well known, is used.

[0020] Here, since the eyes, nose, mouth, etc. have  
15 considerable edges so that these are effectively utilized as the amount of features. Moreover, in the case when the amount of features other than edges is utilized, another image processing means is added. Moreover, in the case when only the image gray-scale value is used as the amount of features,  
20 the image processing section is not required. Moreover, the image processing is not necessarily carried out on the entire image, and it may be carried out for each of collation local areas.

[0021] As illustrated in Fig. 4, the position detection  
25 section 3 is provided with a model generation section 31 for

generating area models of a face, a model storage section 32 for storing the area models generated by the model generation section 31, a position specifying section 33 for specifying the position at which the area model stored in the model storage section 32 is inserted to the digital gray-scale image and the edge image, stored in the image memory 13, a feature-amount extraction section 34 serving as a judgment element acquiring means for extracting the amount of features in each of the judgment element acquiring areas within the area model at the position specified by the position specifying section 33 and a judgment section 36 serving as a judgment means for making a judgment as to whether or not the image in question is a face image by using judgment-use information stored in a judgment-use storage section 35 based upon the amount of features within each of the judgment element acquiring area extracted by the feature-amount extraction section 34.

[0022] As illustrated in Fig. 5, the model generation section 31 generates the area models which are a plurality of preset judgment element acquiring areas corresponding to featuring areas such as the eyes, nose, mouth and cheeks etc. of a face image. In the area models shown in Fig. 5(a), featuring areas are set as areas of the eyes, nose, mouth and cheeks, and judgment element acquiring areas 41 and 42 for the eyes, a judgment element acquiring area 43 for the nose, a judgment element acquiring area 44 for the mouth and a judgment element acquiring

area 45 for the cheek are set. In the area models shown in Fig. 5(b), featuring areas are set as areas of the eyes, mouth and cheeks, and judgment element acquiring areas 41 and 42 for the eyes, a judgment element acquiring area 44 for the mouth and a judgment element acquiring area 45 for the cheek are set.

[0025] Moreover, area models shown in Fig. 6 are referred to as mask models, and each area is represented by a mask. Numeric values 1 and 2 are judgment element acquiring areas for the eyes, numeric value 3 is a judgment element acquiring area for the nose, numeric value 4 is a judgment element acquiring area for the mouth, and numeric value 5 is a judgment element acquiring area for the cheeks. Here, portions having a numeric value of 0 are areas to be excluded. By using such mask expressions, it is possible to generate more detailed area models.

[0032] At the position specified by the position specifying section 33, with respect to the amount of features to be extracted from both of the images, that is, the digital gray-scale image and the edge image, for example, the amount of features represented by the following equations are used. Here, it is supposed that the area model of a face to be utilized is a model shown in Fig. 5(a). Moreover, within the area model,

the judgment element requiring area 41 for the right eye is represented as RE, the judgment element requiring area 42 for the left eye is represented as LE, the judgment element requiring area 43 for the nose is represented as N, the judgment  
5 element requiring area 44 for the mouth is represented as M, the rest of the judgment element requiring areas 45 for the cheeks is represented by C. Furthermore, it is supposed that each of the digital gray-scale image in the image memory 13 and the edge image are images having values of pixels 0 to  
10 255. Moreover,  $P(i)$  is set to be a gray-scale value of a digital gray-scale image at the position  $i$ , and  $E(i)$  is set to be a gray-scale value of an edge image at the position  $i$ .

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-283036

(43) 公開日 平成11年(1999)10月15日

(51) Int.Cl.<sup>5</sup>

G 0 6 T 7/00

識別記号

F I

G 0 6 F 15/62  
15/70

4 6 5 K  
4 5 5 A

審査請求 未請求 請求項の数11 O L (全 10 頁)

(21) 出願番号 特願平10-83736

(22) 出願日 平成10年(1998) 3 月30日

(71) 出願人 000003562

東芝テック株式会社

東京都千代田区神田錦町1丁目1番地

(72) 発明者 羽毛田 卓哉

静岡県三島市南町6番78号 株式会社テック技術研究所内

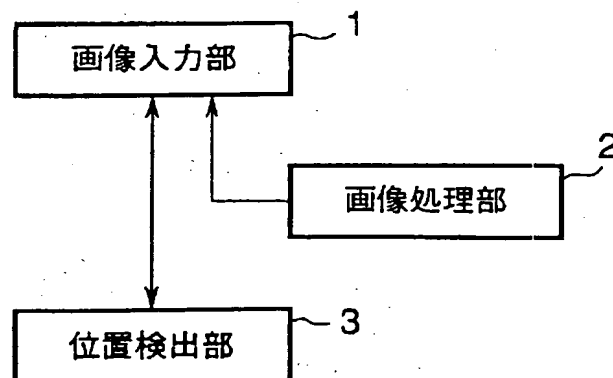
(74) 代理人 弁理士 鈴江 武彦 (外6名)

(54) 【発明の名称】 対象物検出装置及び対象物検出方法

(57) 【要約】

【課題】被識別対象物の特徴領域位置の個体差、サイズの変動、撮影環境の変動を考慮して対象物検出を行って検出率の向上を図り、また、特徴領域の位置関係の検証を不要にして簡潔な処理を実現する。

【解決手段】画像を画像入力部1から顔を含む画像を入力し濃淡画像として記憶し、この濃淡画像を画像処理部2により微分処理してエッジ画像を生成して記憶する。そして、位置検出部3において、濃淡画像とエッジ画像に対して、顔画像の目、鼻、口等の特徴的な領域に対応して複数の判定要素取得領域を設定した領域モデルを使用して、この領域モデルを画像に対して位置を順次指定しつつ当て嵌め、各判定要素取得領域から特徴量を抽出し、この抽出した特徴量に基づいて画像内から顔画像を検出する。



## 【特許請求の範囲】

【請求項 1】 画像を入力する画像入力手段と、検出する被識別対象物画像の特徴的な領域に対応して複数の判定要素取得領域を設定した領域モデルを記憶した記憶手段と、前記画像入力手段により入力した入力画像及びこの入力画像を画像処理して得られる画像の一方又は両方に対して前記記憶手段に記憶した領域モデルを当て嵌める位置を順次指定する位置指定手段と、この位置指定手段にて指定した位置に領域モデルを順次当て嵌める毎に、この領域モデルの各判定要素取得領域から判定要素を取得する判定要素取得手段と、この判定要素取得手段が取得した判定要素に基づいて被識別対象物画像か否かの判定を行う判定手段とからなり、前記判定手段の判定結果により被識別対象物の検出を行うことを特徴とする対象物検出装置。

【請求項 2】 画像を入力する画像入力手段と、検出する顔画像の目、口、鼻、頬等の特徴的な領域に対応して複数の判定要素取得領域を設定した領域モデルを記憶した記憶手段と、前記画像入力手段により入力した入力画像及びこの入力画像を画像処理して得られる画像の一方又は両方に対して前記記憶手段に記憶した領域モデルを当て嵌める位置を順次指定する位置指定手段と、この位置指定手段にて指定した位置に領域モデルを順次当て嵌める毎に、この領域モデルの各判定要素取得領域から判定要素を取得する判定要素取得手段と、この判定要素取得手段が取得した判定要素に基づいて顔画像か否かの判定を行う判定手段とからなり、前記判定手段の判定結果により顔の検出を行うことを特徴とする対象物検出装置。

【請求項 3】 判定要素取得手段は、領域モデルの各判定要素取得領域から判定要素として特徴量を取得することを特徴する請求項 1 又は 2 記載の対象物検出装置。

【請求項 4】 判定要素取得手段は、領域モデルの少なくとも 1 つの判定要素取得領域に対してテンプレート画像を利用して特徴量を取得することを特徴する請求項 3 記載の対象物検出装置。

【請求項 5】 判定要素取得手段は、領域モデルの判定要素取得領域として目、口、鼻の領域の 1 又は複数が設定されたときには、この領域の少なくとも 1 つに対して領域内の濃淡値あるいは濃淡値の微分値を利用して特徴量を取得することを特徴する請求項 3 又は 4 記載の対象物検出装置。

【請求項 6】 判定要素取得手段は、領域モデルの判定要素取得領域として頬の領域が設定されたときには、この領域に対して領域内の濃淡値の分散値を利用して特徴量を取得することを特徴する請求項 3 又は 5 記載の対象物検出装置。

【請求項 7】 判定要素取得手段は、領域モデルの判定要素取得領域として頬の領域が設定されたときには、この領域に対してテンプレート画像を利用して特徴量を取

得することを特徴する請求項 3 又は 5 記載の対象物検出装置。

【請求項 8】 入力した入力画像及びこの入力画像を画像処理して得られる画像の一方又は両方に対して、検出する被識別対象物画像の特徴的な領域に対応して複数の判定要素取得領域を設定した領域モデルを順次位置を指定しながら当て嵌め、領域モデルを当て嵌める毎にこの領域モデルの各判定要素取得領域から判定要素を取得し、この取得した判定要素に基づいて被識別対象物画像か否かの判定を行い、この判定結果により被識別対象物の検出を行うことを特徴とする対象物検出方法。

【請求項 9】 入力した入力画像及びこの入力画像を画像処理して得られる画像の一方又は両方に対して、検出する顔画像の目、口、鼻等の特徴的な領域に対応して複数の判定要素取得領域を設定した領域モデルを順次位置を指定しながら当て嵌め、領域モデルを当て嵌める毎にこの領域モデルの各判定要素取得領域から判定要素を取得し、この取得した判定要素に基づいて顔画像か否かの判定を行い、この判定結果により顔の検出を行うことを特徴とする対象物検出方法。

【請求項 10】 領域モデルの各判定要素取得領域から判定要素として特徴量を取得することを特徴とする請求項 8 又は 9 記載の対象物検出方法。

【請求項 11】 領域モデルの少なくとも 1 つの判定要素取得領域に対してテンプレート画像を利用して特徴量を取得することを特徴とする請求項 10 記載の対象物検出方法。

## 【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、顔や物品など、被識別対象物の検出を行う対象物検出装置及び対象物検出方法に関する。

【0002】

【従来の技術】ある被識別対象物や背景を含む画像から被識別対象物を検出する装置としては、例えば、特開平 7-129770 号公報のものが知られている。これは、例えば、工場における組立て工程や検査工程などにおいて、被識別対象物をテレビカメラで撮影し、この撮影した入力画像と基準となるテンプレート画像との濃淡画像について相関を取ることによって被識別対象物の位置を検出するようにしている。具体的には、探索画像として、テンプレート画像と同じ大きさの局所領域を入力画像中に設定し、これを順にずらしながらテンプレート画像と探索画像との相関値を計算し、相関値が最も高い位置をテンプレートが存在する位置として特定するようになっている。

【0003】また、テンプレートマッチング法を利用して入力画像から顔画像を検出するものとしては、例えば、特開平 9-251534 号公報や特開平 9-44676 号公報が知られている。特開平 9-251534 号



公報のものは、入力画像から顔画像領域を抽出するために、予め登録されている標準顔画像（テンプレート）を全画面にわたって移動させつつ相関値を計算し、最も高い相関値を有する領域を顔領域として抽出するものである。また、特開平9-44676号公報のものは、目を濃淡情報で表したテンプレート画像で顔画像を含む画像を走査し、対象領域の濃淡情報とテンプレート画像の濃淡情報との相関演算を行い、類似度の高い領域を目の候補として抽出する。同様に、鼻を濃淡情報で表したテンプレート画像で顔画像を含む画像を走査して鼻の候補を抽出し、口を濃淡情報で表したテンプレート画像で顔画像を含む画像を走査して口の候補を抽出する。そして、抽出が終了すると、顔領域の抽出を行うが、この時、目、鼻、口の候補の組み合わせについて、予め用意されている目、鼻、口の位置関係と比較検証して画像の中から顔画像を抽出するというものである。

【0004】さらに、顔画像の色情報を利用したものとしては、例えば、特開平9-50528号公報が知られている。これは、入力画像のRGB値から肌色領域を抽出し、この領域に対して自動的にモザイクサイズを決定し、その候補領域をモザイク化し、人物顔辞書と比較して人物顔の有無を判定し人物顔の切り出しを行うというものである。

【0005】

【発明が解決しようとする課題】ところで、顔は個人毎に鼻や目の位置、大きさが様々であり、また、人物の顔の検出の適用場所は工場内のファクトリ・オートメーション装置と違い背景が複雑で、かつ外光環境の変動も大きい場合が多い。また、人物はカメラの前で動きがあるため、人物の前後位置の多少のずれによっても顔サイズは変わり、また、顔の傾きのばらつきも生じる。

【0006】このようなことから、特開平7-129770号公報や特開平9-251534号公報のようなテンプレートマッチング法を利用するものでは、人物の顔を顔全体のテンプレート画像としてそのまま用いることになるため、テンプレート画像に柔軟性がなく、検出率がそれ程高くないという問題があった。すなわち、個人差等を考慮していないので、ある人は検出できるが別の人は検出できないというような事態が生じる。

【0007】また、特開平9-44676号公報のような個々の部品を抽出するような方法では、個々の部品毎に抽出方法を変えたり、候補領域の位置関係を検証したりなど手順が複雑であり、安定した検出が難しい問題があり、特に、背景が複雑な場合においては部品候補の数が大量になる可能性があり、その位置関係を検証することが困難になるという問題があった。さらに、特開平9-50528号公報のような肌色領域を抽出する色情報を用いて予め候補領域を絞り込むような方法では、照明条件の影響を受けやすく安定した検出が困難になる問題があった。

【0008】そこで、請求項1乃至7記載の発明は、被識別対象物の特徴領域の位置の個体差、サイズの変動、撮影環境の変動を考慮して被識別対象物を検出することができて検出率の向上を図ることができ、しかも特徴領域の位置関係を検証する必要がなく簡潔な処理を実現できる対象物検出装置を提供する。また、請求項8乃至11記載の発明は、被識別対象物の特徴領域の位置の個体差、サイズの変動、撮影環境の変動を考慮して被識別対象物を検出することができて検出率の向上を図ることができ、しかも特徴領域の位置関係を検証する必要がなく簡潔な処理を実現できる対象物検出方法を提供する。

【0009】

【課題を解決するための手段】請求項1記載の発明は、画像を入力する画像入力手段と、検出する被識別対象物画像の特徴的な領域に対応して複数の判定要素取得領域を設定した領域モデルを記憶した記憶手段と、画像入力手段により入力した入力画像及びこの入力画像を画像処理して得られる画像の一方又は両方に対して記憶手段に記憶した領域モデルを当て嵌める位置を順次指定する位置指定手段と、この位置指定手段にて指定した位置に領域モデルを順次当て嵌める毎に、この領域モデルの各判定要素取得領域から判定要素を取得する判定要素取得手段と、この判定要素取得手段が取得した判定要素に基づいて被識別対象物画像が否かの判定を行う判定手段とからなり、判定手段の判定結果により被識別対象物の検出を行う対象物検出装置にある。

【0010】請求項2記載の発明は、画像を入力する画像入力手段と、検出する顔画像の目、口、鼻、頬等の特徴的な領域に対応して複数の判定要素取得領域を設定した領域モデルを記憶した記憶手段と、画像入力手段により入力した入力画像及びこの入力画像を画像処理して得られる画像の一方又は両方に対して記憶手段に記憶した領域モデルを当て嵌める位置を順次指定する位置指定手段と、この位置指定手段にて指定した位置に領域モデルを順次当て嵌める毎に、この領域モデルの各判定要素取得領域から判定要素を取得する判定要素取得手段と、この判定要素取得手段が取得した判定要素に基づいて顔画像が否かの判定を行う判定手段とからなり、判定手段の判定結果により顔の検出を行う対象物検出装置にある。

【0011】請求項3記載の発明は、請求項1又は2記載の対象物検出装置において、判定要素取得手段は、領域モデルの各判定要素取得領域から判定要素として特徴量を取得することにある。請求項4記載の発明は、請求項3記載の対象物検出装置において、判定要素取得手段は、領域モデルの少なくとも1つの判定要素取得領域に対してテンプレート画像を利用して特徴量を取得することにある。請求項5記載の発明は、請求項3又は4記載の対象物検出装置において、判定要素取得手段は、領域モデルの判定要素取得領域として目、口、鼻の領域の1又は複数が設定されたときには、この領域の少なくとも

1つに対して領域内の濃淡値あるいは濃淡値の微分値を利用して特徴量を取得することにある。

【0012】請求項6記載の発明は、請求項3又は5記載の対象物検出装置において、判定要素取得手段は、領域モデルの判定要素取得領域として頬の領域が設定されたときには、この領域に対して領域内の濃淡値の分散値を利用して特徴量を取得することにある。請求項7記載の発明は、請求項3又は5記載の対象物検出装置において、判定要素取得手段は、領域モデルの判定要素取得領域として頬の領域が設定されたときには、この領域に対してテンプレート画像を利用して特徴量を取得することにある。

【0013】請求項8記載の発明は、入力した入力画像及びこの入力画像を画像処理して得られる画像の一方又は両方に対して、検出する被識別対象物画像の特徴的な領域に対応して複数の判定要素取得領域を設定した領域モデルを順次位置を指定しながら当て嵌め、領域モデルを当て嵌める毎にこの領域モデルの各判定要素取得領域から判定要素を取得し、この取得した判定要素に基づいて被識別対象物画像か否かの判定を行い、この判定結果により被識別対象物の検出を行う対象物検出方法にある。

【0014】請求項9記載の発明は、入力した入力画像及びこの入力画像を画像処理して得られる画像の一方又は両方に対して、検出する顔画像の目、口、鼻等の特徴的な領域に対応して複数の判定要素取得領域を設定した領域モデルを順次位置を指定しながら当て嵌め、領域モデルを当て嵌める毎にこの領域モデルの各判定要素取得領域から判定要素を取得し、この取得した判定要素に基づいて顔画像か否かの判定を行い、この判定結果により顔の検出を行う対象物検出方法にある。

【0015】請求項10記載の発明は、請求項8又は9記載の対象物検出方法において、領域モデルの各判定要素取得領域から判定要素として特徴量を取得することにある。請求項11記載の発明は、請求項10記載の対象物検出方法において、領域モデルの少なくとも1つの判定要素取得領域に対してテンプレート画像を利用して特徴量を取得することにある。

【0016】

【発明の実施の形態】本発明の実施の形態を図面を参照して説明する。

(第1の実施の形態)なお、この実施の形態は被識別対象物検出として顔検出を例にした場合について述べる。

【0017】図1は対象物検出装置の全体構成を示すブロック図で、人物の顔を含む画像を入力する画像入力手段としての画像入力部1、この画像入力部1により入力した画像に対して所定の処理を施す画像処理部2、前記画像入力部1が入力した画像情報及び前記画像処理部2が処理した画像情報から顔画像の位置を検出する位置検出部3とで構成している。

【0018】前記画像入力部1は、図2に示すように、人物の顔を撮影し、顔を含むデジタル濃淡画像情報を出力するCCDカメラ11と、このCCDカメラ11からのデジタル濃淡画像情報を取込む画像入力ボード12と、この画像入力ボード12が取込んだデジタル濃淡画像情報を記憶する画像メモリ13とで構成している。なお、入力する画像はカラー画像でもよい。

【0019】前記画像処理部2は、前記画像メモリ13からデジタル濃淡画像情報を読み出し、これを微分処理してエッジ画像情報に変換し、前記画像メモリ13に記憶するようになっている。すなわち、図3の(a)に示すような濃淡画像を微分処理して図3の(b)に示すようなエッジ画像を生成し画像メモリ13に記憶するようになっている。エッジの抽出には、よく知られているラプラシアンフィルタやSobelフィルタなどを使用する。

【0020】なお、目、鼻、口などはかなりのエッジが存在するので、特徴量として利用するのは有効である。また、エッジ以外の特徴量を利用する場合は他の画像処理手段を追加すればよい。また、画像濃淡値のみを特徴量として用いる場合は画像処理部はなくてもよい。さらに、画像処理は画像全体で行わずに被照合局所領域毎に行ってもよい。

【0021】前記位置検出部3は、図4に示すように、顔の領域モデルを生成するモデル生成部31と、このモデル生成部31が生成した領域モデルを記憶するモデル記憶部32と、このモデル記憶部32に記憶した領域モデルを前記画像メモリ13に記憶したデジタル濃淡画像及びエッジ画像に当て嵌める位置を指定する位置指定部33と、この位置指定部33が指定した位置における領域モデル内の複数の判定要素取得領域の特徴量を抽出する判定要素取得手段としての特徴量抽出部34と、この特徴量抽出部34が抽出した各判定要素取得領域の特徴量から判定用記憶部35に記憶している判定用情報を使用して顔画像か否かの判定を行う判定手段としての判定部36とで構成している。

【0022】前記モデル生成部31は、図5に示すように顔画像の目、鼻、口、頬などの特徴的な領域に対応して複数の判定要素取得領域を設定した領域モデルを生成するようになっている。図5の(a)の領域モデルは、特徴的な領域を目、鼻、口、頬の領域とし、目に対する判定要素取得領域41、42、鼻に対する判定要素取得領域43、口に対する判定要素取得領域44、残りの頬に対する判定要素取得領域45を設定したものである。図5の(b)の領域モデルは、特徴的な領域を目、口、頬の領域とし、目に対する判定要素取得領域41、42、口に対する判定要素取得領域44、頬に対する判定要素取得領域45を設定したものである。

【0023】図5の(c)の領域モデルは、特徴的な領域を目、頬の領域とし、目に対する判定要素取得領域41、42、頬に対する判定要素取得領域45を設定した

ものである。図 5 の(d) の領域モデルは、特徴的な領域を目、鼻、頬の領域とし、目に対する判定要素取得領域 4 1、4 2、鼻に対する判定要素取得領域 4 3、頬に対する判定要素取得領域 4 5 を設定したものである。

【0024】図 5 の(e) の領域モデルは、特徴的な領域を目、鼻、口、頬の領域とし、左右の目を 1 つに設定した判定要素取得領域 4 6、鼻に対する判定要素取得領域 4 3、口に対する判定要素取得領域 4 4、頬に対する判定要素取得領域 4 5 を設定したものである。図 5 の(f) の領域モデルは、特徴的な領域を目、鼻、口、頬、髪領域とし、目に対する判定要素取得領域 4 1、4 2、鼻に対する判定要素取得領域 4 3、口に対する判定要素取得領域 4 4、頬に対する判定要素取得領域 4 5、髪に対する判定要素取得領域 4 7 を設定したものである。これらの領域モデルはウィンドウモデルと呼ばれるものである。

【0025】また、図 6 に示す領域モデルはマスクモデルと呼ばれるもので、各領域をマスクで表現するようになっている。数値 1 及び 2 は目に対する判定要素取得領域であり、数値 3 は鼻に対する判定要素取得領域であり、数値 4 は口に対する判定要素取得領域であり、数値 5 は頬に対する判定要素取得領域である。なお、数値 0 の部分は除外する領域である。このようなマスク表現により、より詳細な領域モデルを生成することができる。

【0026】どの領域モデルを利用した場合でも、1 つのモデルである程度の顔のサイズの違いや顔部品の位置の個人差を吸収することはできるが、より多種の顔のサイズに対応させるためにはサイズの異なった領域モデルを利用すればよい。なお、予め必要な領域モデルを作成してモデル記憶部 3 2 に記憶しておけばモデル生成部 3 1 は省略できる。

【0027】前記モデル記憶部 3 2 は、図 5 に示すような顔の領域モデルを記憶する。ウィンドウモデルの場合、ウィンドウの左上を原点として、各判定要素取得領域を表す各ウィンドウの座標や幅、高さ等を記憶する。マスクモデルの場合、図 6 に示すマスク値をそのまま記憶する。なお、モデルの記憶方法としてはその他様々な方法があり、これらに限定するものではない。

【0028】前記モデル記憶部 3 2 に記憶した領域モデルを、前記画像入力部 1 からのデジタル濃淡画像及び前記画像処理部 2 で処理したエッジ画像に当て嵌めて顔画像の検出を行う。このような領域モデル M を当て嵌めることにより、図 7 に示すような様々な人物、また、多少顔を傾けた人、眼鏡をかけた人の顔の検出が可能になる。なお、計算機プログラム上では顔のモデルをメモリに記憶するという形を取らなくてもパラメータとして領域モデル内の各領域を指定することは可能であり、このような場合もプログラムとして記憶していると見なす。

【0029】前記位置指定部 3 3 及び特徴量抽出部 3 4 は、前記画像メモリ 1 3 に記憶したデジタル濃淡画像及

びエッジ画像の両画像に対して図 8 に示すように領域モデル M を当て嵌める位置を指定し、その位置における 2 つの画像から領域モデル M 内の各判定要素取得領域の特徴量を抽出する。すなわち、前記位置指定部 3 3 は、基本的には 2 つの画像全体に順次領域モデルの位置を指定し、前記特徴量抽出部 3 4 は、指定された各位置において領域モデル M 内の各判定要素取得領域の特徴量を抽出する。そして、領域モデル M が M' の位置に指定されたときに最も顔らしい特徴量が抽出されることになる。

【0030】なお、その他の特徴量を用いる場合は、画像処理部 2 においてその他の特徴量を含む画像を入力画像から生成して画像メモリ 1 3 に記憶し、この記憶した画像を利用すればよい。ここで入力画像がカラー画像の場合は、このカラー入力画像から肌色らしい領域を抽出し、その領域に対してのみ領域モデルの位置を指定してもよい。

【0031】このように前処理により検出領域を絞ってから領域モデルの位置を指定することもできる。また、カラー画像を利用して肌色という特徴を用いることも可能である。なお、ここでは領域モデルの当て嵌め方として入力画像に対して領域モデル全体を指定位置を移動させながら順次当て嵌めたが、必ずしもこれに限定するものではなく、領域モデルの各判定要素取得領域をばらばらにして 1 つずつ指定位置を移動させながら順次当て嵌めてもよい。

【0032】前記位置指定部 3 3 で指定した位置において、デジタル濃淡画像及びエッジ画像の両画像から抽出する特徴量としては、例えば、次式で表現される特徴量を用いる。なお、利用する顔の領域モデルは図 5 の(a) に示すモデルとする。また、領域モデル内の右目の判定要素取得領域 4 1 を R E、左目の判定要素取得領域 4 2 を L E、鼻の判定要素取得領域 4 3 を N、口の判定要素取得領域 4 4 を M、残りの頬の判定要素取得領域 4 5 を C とする。また、前記画像メモリ 1 3 中のデジタル濃淡画像、エッジ画像ともに各画素 0 ~ 2 5 5 の値をとる画像とする。さらに、 $P(i)$  を位置  $i$  におけるデジタル濃淡画像の濃淡値、 $E(i)$  を位置  $i$  におけるエッジ画像の濃淡値とする。

【0033】右目の判定要素取得領域 R E の低輝度特徴量の合計は、

$$\text{【数 1】} \quad \text{RE\_d (Right Eye darkness)} = \sum_{i \in \text{Area RE } (P(i) <= 100)} (100 - P(i))$$

【0034】右目の判定要素取得領域 R E のエッジ特徴量の合計は、

$$\text{【数 2】} \quad \text{RE\_e (Right Eye Edge)} = \sum_{i \in \text{Area RE}} E(i)$$

【0035】左目の判定要素取得領域 L E の低輝度特徴量の合計は、

【数 3】

$$LE\_d (\text{Left Eye darkness}) = \sum_{i \in \text{Area LE} (P(i) <= 100)} (100 - P(i))$$

【0036】左目の判定要素取得領域LEのエッジ特徴量の合計は、

【数 4】

$$LE\_e (\text{Left Eye Edge}) = \sum_{i \in \text{Area LE}} E(i)$$

【0037】鼻の判定要素取得領域Nの低輝度特徴量の合計は、

【数 5】

$$N\_d (\text{Nose darkness}) = \sum_{i \in \text{Area N} (P(i) <= 100)} (100 - P(i))$$

【0038】鼻の判定要素取得領域Nのエッジ特徴量の合計は、

$$C\_a (\text{Cheak average}) = \frac{1}{n} \sum_{i \in \text{Area C}} P(i) \quad (n \text{ は Area C の全要素数})$$

【0042】頬の判定要素取得領域Cの輝度値分散は、

$$C\_v (\text{Cheak variance}) = \frac{1}{n} \sum_{i \in \text{Area C}} (P(i) - C\_a)^2 \quad (n \text{ は Area C の全要素数})$$

【0043】となる。なお、頬の判定要素取得領域Cの輝度値分散については、目、鼻、口を取り除いた残りの頬領域Cは輝度分散が小さいという特徴を持っている。

【0044】判定のために上記各式で表される各特徴量を抽出し、メモリに記憶する。こうして抽出された特徴量の例を示すと図9に示すようになる。各判定要素取得領域の特徴量は人物により様々である。すなわち、個人差がある。撮影環境の違いや顔サイズの多少の違い、顔の傾きの変動等もこのデータ中に含まれる。位置指定部33で指定され、特徴量抽出部34で抽出された領域モデルの各領域の特徴量より、判定部36において指定された領域モデルの位置が顔であるか否かを判定する。

【0045】判定には、例えばファジーメンバーシップ関数を利用する。例えば、右目の低輝度画素量としての正しさを $\mu_{RE-d}$ として、ファジーメンバーシップ関数を図10のように定義する。この関数はある指定された位置における領域モデル内の右目の判定要素取得領域REの低輝度画素量RE-dを入力して、その値に応じて0.0～1.0の値を出力する。

【0046】図10ではRE-dが1, 200～10, 00

$$F = \mu_{RE-d} \times \mu_{RE-e} \times \mu_{LE-d} \times \mu_{LE-e} \times \mu_{N-d} \times \mu_{N-e} \\ \times \mu_{M-d} \times \mu_{M-e} \times \mu_{C-a} \times \mu_{C-v}$$

【0049】のような関数を利用する。この式の場合、個々のメンバーシップ関数の出力が1.0に近いほどFは1.0に近くなる。但し、この式の場合、全てが1.0でないとFは1.0にならない。

【0050】Fの式については様々な方法が考えられ、ファジールールを利用したり、代数和で求めることもできる。このFの値が最も大きな部分を顔の位置とし、入

【数 6】

$$N\_e (\text{Nose edge}) = \sum_{i \in \text{Area N}} E(i)$$

【0039】口の判定要素取得領域Mの低輝度特徴量の合計は、

【数 7】

$$M\_d (\text{Mouse darkness}) = \sum_{i \in \text{Area M} (P(i) <= 100)} (100 - P(i))$$

【0040】口の判定要素取得領域Mのエッジ特徴量の合計は、

【数 8】

$$M\_e (\text{Mouse edge}) = \sum_{i \in \text{Area M}} E(i)$$

【0041】頬の判定要素取得領域Cの輝度値平均は、

【数 9】

【数 10】

0と計測された場合、 $\mu_{RE-d}$ は1.0を出力する。この範囲については、図9のようなサンプル画像から得られた特徴量の個人差、変動を参考に決定する。その他の範囲の値の場合は、台形の左右の辺をなす一次関数の出力が $\mu_{RE-d}$ の値となる。 $\mu_{RE-d}$ が1.0に近いほど右目の低輝度画素量として正しいことになる。

【0047】この台形のメンバーシップ関数の決定は、図9に示したように予め様々な学習用画像から顔の特徴量を計測し、この計測した特徴量に基づいて決定する。同様に、図9に示したような各判定要素取得領域のサンプル画像の特徴量から、RE-e、LE-d、LE-e、N-d、N-e、M-d、M-e、C-a、C-vに対するファジーメンバーシップ関数を定義し、それぞれの出力 $\mu_{RE-d}$ 、 $\mu_{RE-e}$ 、 $\mu_{LE-d}$ 、 $\mu_{LE-e}$ 、 $\mu_{N-d}$ 、 $\mu_{N-e}$ 、 $\mu_{M-d}$ 、 $\mu_{M-e}$ 、 $\mu_{C-a}$ 、 $\mu_{C-v}$ を算出する。

【0048】それぞれのメンバーシップ関数は、前記判定用記憶部35に記憶しておけばよい。そして、最終的にその位置が顔であるか否かを判定する。顔らしさFとしては、例えば、

【数 11】

力画像上のその位置に円を描くと、例えば図11のような画像が得られる。以上の判定に関しては、必ずしもファジー理論を利用する必要はない。また、より確実な検出を行う場合には、検出された位置において、さらに本当の顔であるかの検証を行う。

【0051】このように、被識別対象物である顔を、目、鼻、口、頬等の特徴的な判定要素取得領域に分けて

表現した顔の領域モデルを作成し、各判定要素取得領域の位置関係を維持したこの領域モデルを入力画像及び入力画像を画像処理した画像の両方に当て嵌める位置を指定し、その位置における領域モデル内の各判定要素取得領域の特徴量を計測し、領域モデル全体で顔か否かの判定を行っているので、個々の顔の目、鼻等の部品の位置関係を検証する必要がなく、簡潔な処理が実現できる。また、顔の個人差による部品の位置ずれや大きさの違いなどを吸収し、かつ、被写人物の前後位置における顔のサイズの変動や多少の顔の傾きも吸収できるため、簡潔で安定した顔の検出ができる。

【0052】なお、この実施の形態では、領域モデルを入力画像及びこの入力画像を画像処理した画像の両方に当て嵌めて各判定要素取得領域の特徴量を計測して顔の判定を行ったが必ずしもこれに限定するものではなく、領域モデルを入力画像及びこの入力画像を画像処理した画像のいずれか一方に当て嵌めて各判定要素取得領域の特徴量を計測して顔の判定を行ってもよい。

【0053】(第2の実施の形態) この実施の形態も顔検出に関し、基本的には第1の実施の形態と同様であり、異なる点は、領域モデルにおける頬の判定要素取得領域をテンプレート画像に置き換えた点である。

【0054】すなわち、図12の(a)に示すように、マスク表現した顔の領域モデル51における頬の判定要素取得領域をテンプレート画像52に置き換える。このテンプレート画像は図12の(b)に示すような構成になっている。なお、数値1の領域は右目の判定要素取得領域53であり、数値2の領域は左目の判定要素取得領域54であり、数値3の領域は鼻の判定要素取得領域55であり、数値4の領域は口の判定要素取得領域56である。このテンプレート画像52は、実際の顔画像から切り取った濃淡画像でもよい。例えば、頬はほぼ平面なので濃淡値1,1,1,1,1,1,...のような濃淡画像を使用する。これらの画像とテンプレートマッチングしたときの頬の領域の類似度C-similarを求める。

【0055】この頬領域のテンプレート画像をgとし、入力濃淡画像から指定された位置の頬領域をgと同じ大きさで切り出した画像をfとすると、その類似度C-sは、

【数12】

$$C_s = \frac{f \cdot g}{\|f\| \|g\|}$$

【0056】で表される。類似度C-sは0～1の間の値となる。この類似度を第1の実施の形態と同様に様々な人物より取得して学習し、判定用データとしてメモリに記憶し、図13に示すように、他の領域の特徴量と一緒に利用する。実際の顔の判定は類似度を特徴量の1つと考え、第1の実施の形態と同様な処理を行えばよい。

【0057】このように顔全体のテンプレートマッチン

グではなく、領域モデルのうち、頬の領域のテンプレートマッチングと他の領域について特徴量を組み合わせても第1の実施の形態の場合と同様の顔検出ができる。なお、この実施の形態では頬領域についてテンプレートとの類似度を求めるようにしたが必ずしもこれに限定するものではなく、テンプレート画像との距離等を用いてもよい。その他、類似度判定方法としては様々な方法があり、いずれも適用できる。

【0058】なお、前述した各実施の形態では、入力画像の画像全体に対して、領域モデルを当て嵌める位置を順次指定して特徴量を抽出し顔検出を行うようにしたが必ずしもこれに限定するものではなく、図14に示すように、予め背景が既知の場合には、画像全体からその既知の部分57を取り除いて候補領域58を生成し、その候補領域58に対し、前述した各実施の形態と同様に領域モデルを当て嵌める位置を順次指定して特徴量を抽出し顔検出を行ってもよい。このようにすれば特徴量の抽出処理がより迅速になる。

【0059】(第3の実施の形態) この実施の形態は、被識別対象物検出として顔以外の物品検出に適用した例について述べる。具体的には、シーンの中から一方通行の道路標識を検出例について述べる。カラー画像を用いて色特徴を利用してもよいが、ここでは濃淡画像のみを利用する。対象は正面、水平に配置されたシーン画像中の一方通行の標識とする。領域モデルとして、図15に示すような2つの領域61, 62からなる領域モデル60を使用する。

【0060】矢印部の領域61をA1、その他の領域62をA2とする。次に各領域A1、A2において利用する特徴量を決定する。一般に、一方通行の標識においては領域A1の輝度はその他の領域A2に比べて高く、また、領域A1、A2ともに輝度の分散値は非常に小さい。そこで、2つの領域A1、A2とも特徴量として平均輝度と輝度分散値を利用すると、それぞれA1-a、A1-v、A2-a、A2-vとする。そして、様々な一方通行の道路標識のサンプル画像からそれぞれの特徴量を計測し、図16に示すような計測値を得る。そして、この学習結果から各領域A1、A2の正しさとしては、例えば、図17に示すファジーメンバーシップ関数を定義する。ここでは、領域A1の平均輝度の正しさを $\mu_{A1-a}$ として定義している。同様に、領域A1の輝度分散値の正しさを $\mu_{A1-v}$ として定義し、領域A2の平均輝度の正しさを $\mu_{A2-a}$ として定義し、領域A2の輝度分散値の正しさを $\mu_{A2-v}$ として定義している。

【0061】実際の検出においては、まず、標識65を含む図18のようなシーンの画像を入力し、シーン画像上を領域モデル60を順次移動させ、その都度各領域A1、A2の平均輝度や輝度分散値を計測し、例えば、 $F = \mu_{A1-a} \times \mu_{A1-v} \times \mu_{A2-a} \times \mu_{A2-v}$ のような計算式においてその場所に標識65がある確からしさFを計算し、

ある閾値よりもFが大きい位置に標識65があると検出する。このように、被識別対象物である標識65をその標識画像の特徴的な判定要素取得領域A1、A2に分けて表現した顔の領域モデル60を作成し、各判定要素取得領域の位置関係を維持したこの領域モデル60をシーン画像上に当て嵌める位置を指定しつつ領域モデル内の各判定要素取得領域の特徴量を計測し、領域モデル全体で標識か否かの判定を行っているので、簡潔な処理が実現でき、簡潔で安定した標識の検出ができる。

#### 【0062】

【発明の効果】請求項1乃至7記載の発明によれば、被識別対象物の特徴領域の位置の個体差、サイズの変動、撮影環境の変動を考慮して被識別対象物を検出することができて検出率の向上を図ることができ、しかも特徴領域の位置関係を検証する必要がなく簡潔な処理を実現できる対象物検出装置を提供できる。

【0063】また、請求項2乃至7記載の発明によれば、特に顔を含む画像に対して、目、鼻、口、頬等の特徴的な領域に分けて表現した顔の領域モデルを作成し、領域の位置関係を維持したこの領域モデルを画像に当て嵌める位置を順次指定して領域モデルの各判定要素取得領域から特徴量などの判定要素を取得して顔の判定を行うので、個々の顔の目、鼻、口等の部品の位置関係を検証する必要がなく、簡潔な処理ができ、また、顔の個人差による部品の位置ずれや大きさの違いなどを吸収でき、かつ被写人物の前後位置における顔のサイズの変動や多少の顔の傾きも吸収でき、簡潔で安定した顔の検出ができる対象物検出装置を提供できる。

【0064】また、請求項8乃至11記載の発明によれば、被識別対象物の特徴領域の位置の個体差、サイズの変動、撮影環境の変動を考慮して被識別対象物を検出することができて検出率の向上を図ることができ、しかも特徴領域の位置関係を検証する必要がなく簡潔な処理を実現できる対象物検出方法を提供できる。

【0065】また、請求項9乃至11記載の発明によれば、特に顔を含む画像に対して、目、鼻、口、頬等の特徴的な領域に分けて表現した顔の領域モデルを作成し、領域の位置関係を維持したこの領域モデルを画像に当て嵌める位置を順次指定して領域モデルの各判定要素取得領域から特徴量などの判定要素を取得して顔の判定を行うので、個々の顔の目、鼻、口等の部品の位置関係を検証する必要がなく、簡潔な処理ができ、また、顔の個人差による部品の位置ずれや大きさの違いなどを吸収でき、かつ被写人物の前後位置における顔のサイズの変動や多少の顔の傾きも吸収でき、簡潔で安定した顔の検出ができる対象物検出方法を提供できる。

#### 【図面の簡単な説明】

【図1】本発明の第1の実施の形態を示す全体構成のブロック図。

【図2】同実施の形態における画像入力部の構成を示すブロック図。

【図3】同実施の形態において入力した濃淡画像と画像処理したエッジ画像を示す図。

【図4】同実施の形態における位置検出部の構成を示すブロック図。

【図5】同実施の形態において使用可能なウィンドウモデル型の領域モデルの例を示す図。

【図6】同実施の形態において使用可能なマスクモデル型の領域モデルの例を示す図。

【図7】同実施の形態において領域モデルを顔に当て嵌める各種例を示す図。

【図8】同実施の形態において領域モデルを当て嵌める位置指定と特徴量抽出を説明するための図。

【図9】同実施の形態においてサンプル顔画像から抽出した各領域の特徴量の例を示す図。

【図10】同実施の形態における右目の判定要素取得領域の低輝度画素量RE-dのファジーメンバーシップ関数例を示す図。

【図11】同実施の形態における顔の位置検出結果例を示す図。

【図12】本発明の第2の実施の形態を示すもので、マスク表現した顔の領域モデルにおける頬の判定要素取得領域へのテンプレート画像の利用を示す図。

【図13】同実施の形態においてサンプル顔画像から抽出した各領域の特徴量の例を示す図。

【図14】領域モデルを当て嵌める位置指定と特徴量抽出の他の例を説明するための図。

【図15】本発明の第3の実施の形態における領域モデルの例を示す図。

【図16】同実施の形態においてサンプル画像から抽出した各領域の特徴量の例を示す図。

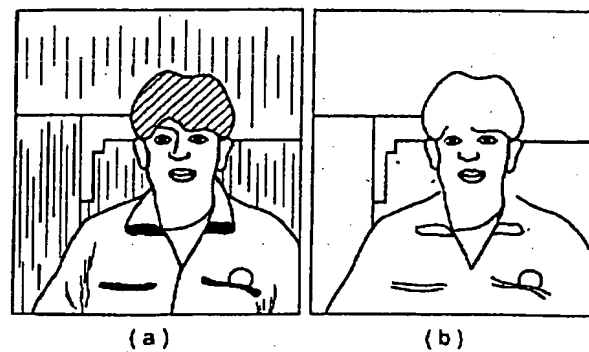
【図17】同実施の形態における領域A1の平均輝度A1-aのファジーメンバーシップ関数例を示す図。

【図18】同実施の形態におけるシーン画像例を示す図。

#### 【符号の説明】

- 1…画像入力部
- 2…画像処理部
- 3…位置検出部
- 13…画像メモリ
- 32…モデル記憶部
- 33…位置指定部
- 34…特徴量抽出部
- 36…判定部

【図 3】



{ b }



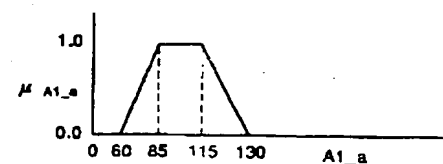
【図 8】



【☒ 1 1】



【圖 17】



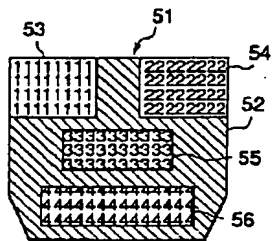
【図 9】

	RE_darkness	RE_edge	LE_darkness	LE_edge	N_darkness	N_edge	M_darkness	.....
人物A	5352	10288	4694	9701	2970	9339	8020	.....
人物B	4278	11514	5200	10009	2865	9491	5855	.....
人物C	6108	9506	8028	7975	1404	7517	6870	.....
人物D	4598	10194	6921	8266	2034	8301	6428	.....
人物E	4651	10756	7226	8247	1904	8047	6550	.....
人物F	1980	10670	1153	11169	1437	10453	2803	.....
人物G	2296	9085	1820	8482	1752	10810	2955	.....
人物H	3213	8139	3001	7800	1715	9714	3503	.....
人物I	4529	8244	3847	7519	1675	8908	3544	.....
人物J	2744	9630	3002	8418	2298	10717	4312	.....
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

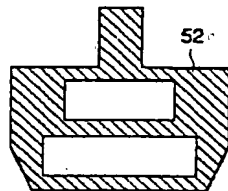
【図 14】



【図 12】



(a)



(b)

【図 13】

RE_edge	LE_darkness	LE_edge	N_darkness	N_edge	M_darkness	....	C_similar
10288	4694	9701	2970	9339	8020	....	0.98
11514	5200	10009	2865	9491	5855	....	0.97
9505	8028	7975	1404	7517	6870	....	0.97
10194	6921	8266	2034	8301	6428	....	0.96
10756	7226	8247	1904	8047	6550	....	0.98
10670	1153	11169	1437	10453	2803	....	0.99
9085	1820	8482	1752	10810	2955	....	0.95
8139	3001	7800	1715	9714	3503	....	0.98
8244	3847	7519	1675	8908	3544	....	0.99
9630	3002	8418	2298	10717	4312	....	0.96

【図 16】

A1 a	A1 v	A2 a	A2 v
100	5	200	4
110	4	187	2
108	2	198	6
112	6	179	3
109	4	211	4
119	7	197	7
110	3	180	5
102	1	186	4
120	8	193	3
116	5	188	6
⋮	⋮	⋮	⋮

【図 18】

